DOCUMENT RESUME

ED 189 586 CS 005 570

TITLE	Testing and Evaluation in Reading and Communication Skills: Abstracts of Doctoral Dissertations Published in "Dissertation Abstracts International," January through June 1980 (Vol. 40 Nos. 7 through 12).
INSTITUTION	ERIC Clearinghouse on Reading and Communication Skills, Urbana, Ill.
PUB DATE	80
NOTE	20p.
AVAILABLE FROM	Copies of the dissertations are available from University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106.
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Annotated Bibliographies; Basic Skills; Cloze Procedure; *Communication Skills; *Doctoral Dissertations; *Educational Research; Elementary Secondary Education; Higher Education; Oral Reading; Reading Difficulties; *Reading Skills; *Reading Tests; Student Attitudes; Teaching Methods; Test Construction; *Testing; Test Validity; Writing Skills

ABSTRACT

This collection of abstracts is part of a continuing series providing information on recent doctoral dissertations. The 35 titles deal with a variety of topics, including the following: (1) cloze procedure; (2) the monitoring and facilitation of reading growth; (3) the effect of immediate item feedback on the reliability and validity of verbal ability test scores; (4) performance variables in essay testing; (5) a process model for identifying and selecting proficiency standards in the basic skills of reading, writing, and mathematics; (6) schema theory as a theoretical framework for the cloze procedure; (7) the verification of the test of affixes in syntactic structures; (8) techniques for the identification of children with reading disabilities; (9) an instrument to measure elementary school teachers' knowledge of a subskills approach to teaching reading; (10) predicting success on a minimum competency examination; (11) comparison of various standardized reading tests; (12) measuring the difficulty of technical writing; (13) two modes of analyzing teacher and student attitudes toward writing; (14) teachers' comments on students' writing; (15) information gain as a measure of reading comprehension; (16) mothers' estimates of their children's reading achievement; and (17) the attitudes of senior high school students toward student-evaluated composition. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED189586

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Testing and Evaluation in Reading and Communication Skills:

Abstracts of Doctoral Dissertations Published in Dissertation Abstracts International, January through June 1980 (Vol. 40 Nos. 7 through 12)

Compiled by the staff of the

ERIC Clearinghouse on Reading and Communication Skills

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

University Microfilms International

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS205570

The dissertation titles contained here are published with permission of the University Microfilms International, publishers of Dissertation Abstracts International (copyright © 1980 by University Microfilms International), and may not be reproduced without their prior permission.

This bibliography has been compiled as part of a continuing series designed to make information on relevant dissertations available to users of the ERIC system. Monthly issues of Dissertation Abstracts International are reviewed in order to compile abstracts of dissertations on related topics, which thus become accessible in searches of the ERIC data base. Ordering information for the dissertations themselves is included at the end of the bibliography.

Abstracts of the following dissertations are included in this collection:

Alderson, John Charles
A STUDY OF THE CLOZE PROCEDURE WITH NATIVE AND NON-NATIVE SPEAKERS OF ENGLISH

Arreola, Mona Jean
THE DIFFERENTIAL EFFECTS OF SYNONYMIC AND EXACT-WORD SCORING OF CLOZE TESTS UPON THE IDENTIFICATION OF INDEPENDENT, INSTRUCTIONAL, AND FRUSTRATION READING LEVELS

Austin, Cecil
MONITORING AND FACILITATION OF READING GROWTH

Benson, Gordon Guy Edward
THE EFFECT OF IMMEDIATE ITEM FEEDBACK ON THE RELIABILITY AND VALIDITY OF VERBAL ABILITY TEST SCORES

Biola, Heather Roberts
PERFORMANCE VARIABLES IN ESSAY TESTING

Bucher, Louis David
A PROCESS MODEL FOR IDENTIFYING AND SELECTING PROFICIENCY STANDARDS IN THE BASIC SKILLS OF READING, WRITING, AND MATHEMATICS

Cothran, Ann Hollingsworth
A STUDY OF THE CLOZE PROCEDURE UNDER THREE DIFFERENT CONDITIONS AT TWO COGNITIVE LEVELS

Davenport, Betty Munch
SCHEMA THEORY AS A THEORETICAL FRAMEWORK FOR THE CLOZE PROCEDURE

Dogger, Barbara Teresa
THE VERIFICATION OF THE TEST OF AFFIXES IN SYNTACTIC STRUCTURES: A STUDY OF DERIVATIONAL MORPHOLOGY AS A LANGUAGE CORRELATE FOR COLLEGE-LEVEL READING PROFICIENCY

Dominick Fields, Nancy Anne
A COMPARISON OF TECHNIQUES FOR THE IDENTIFICATION OF CHILDREN WITH READING DISABILITIES

Dryden, Beverly Joan
THE DEVELOPMENT OF A DIAGNOSTIC READING ATTITUDE ASSESSMENT FOR SECOND GRADE

Fedorko, Frederick John
A COMPARISON OF QUANTITATIVE AND QUALITATIVE METHODS OF ANALYZING THE ORAL READING BEHAVIOR OF FOURTH-GRADE STUDENTS

Harper, Betty Jean
THE CONSTRUCTION OF AN INSTRUMENT TO MEASURE ELEMENTARY TEACHERS' KNOWLEDGE OF A SUBSKILLS APPROACH TO TEACHING READING

Henderson, Donald Joseph
PREDICTING SUCCESS ON A MINIMUM COMPETENCY EXAMINATION

Herzog, Katharine Dora
A COMPREHENSION STUDY OF THREE VERSIONS OF THE GOSPEL OF MARK AND THEIR APPROPRIATENESS FOR USE BY SIXTH GRADERS WITH VARIOUS READING ABILITIES

Hodges, Jane Clayton
A COMPARISON OF THE READING SUBTEST OF THE CALIFORNIA ACHIEVEMENT TESTS, THE GRAY ORAL READING TEST, AND THE PUBLISHERS' PLACEMENT TESTS FOR THE STATE-ADOPTED BASAL READING SERIES

Hull, Leon Clinton, Jr.
BEYOND READABILITY: MEASURING THE DIFFICULTY OF TECHNICAL WRITING

Jorgenson, Jeffrey Lee
A STUDY OF THE DIFFERENCES BETWEEN READING SCORES: AS ACHIEVED ON SILENT INFORMAL READING INVENTORIES AND ORAL INFORMAL READING INVENTORIES BY ELEMENTARY STUDENTS

King, Barbara Lynn
TWO MODES OF ANALYZING TEACHER AND STUDENT ATTITUDES TOWARD WRITING: THE EMIG ATTITUDE SCALE AND THE KING CONSTRUCT SCALE

King, Jean Anne
TEACHERS' COMMENTS ON STUDENTS' WRITING: A CONCEPTUAL ANALYSIS AND EMPIRICAL STUDY

Kirby, Timothy Stephen
A COMPARATIVE STUDY OF SCREENING TESTS IN THE EARLY IDENTIFICATION OF DYSLEXIA

Kyhos, Ronald Milton
AN ANALYSIS OF THE ATTITUDES OF SENIOR HIGH SCHOOL STUDENTS TOWARD STUDENT-EVALUATED COMPOSITION

Lombardo, Maria Elizabeth
THE CONSTRUCTION AND VALIDATION OF THE LISTENING AND READING COMPONENTS OF THE ENGLISH AS A SECOND LANGUAGE ASSESSMENT BATTERY

Lubaus, Zane
AN EXPLORATORY STUDY OF DIFFERENCES BETWEEN TWO ENGLISH READING ACHIEVEMENT TESTS WITH HISPANIC BILINGUAL STUDENTS AND MONOLINGUAL ENGLISH STUDENTS

Miller, Henrietta Clay
THE RELATIONSHIP BETWEEN ABILITY GROUP PLACEMENT AND TWO STANDARDIZED READING MEASURES, A CRITERION-REFERENCED READING MEASURE, AND A MEASURE OF PARADIGMATIC LANGUAGE AMONG SELECTED FOURTH GRADE SUBJECTS

Miura, Steven Yasuo
AN EXAMINATION OF RELATIONAL CONTROL CODING INSTRUMENTATION: CONCEPTUAL, CONSTRUCT, AND PREDICTIVE VALIDITY

Ozier, Paul William
THE CONCURRENT VALIDITY OF CLEP COMPOSITION SCORES IN RELATION TO WRITING PERFORMANCE

Perry, Leslie Anne Devaney
A COMPARATIVE STUDY OF THE PREDICTIVE VALIDITY OF TWO READING READINESS TESTS

Reap, Mary
THE USE OF CLOZE PROCEDURE IN THE MEASUREMENT OF SELECTED ENGLISH SYNTACTIC STRUCTURES AND ITS RELATIONSHIP TO COGNITIVE STYLE

Reynolds, Marianne Clifford
INFORMATION GAIN AS A MEASURE OF READING COMPREHENSION

Richardson, Nelda Savacool
THE ORFF TECHNIQUE AS AN INFORMAL DIAGNOSTIC SCREENING FOR READING READINESS

Sullivan, Vincent Joseph
A STUDY TO DETERMINE THE OPTIMUM READABILITY LEVEL OF A TEXT IN RELATION TO STUDENTS' READING LEVELS AND ACHIEVEMENT IN A COMMUNITY COLLEGE INTRODUCTORY PSYCHOLOGY COURSE

Trunnel, Vesta Mickel
MOTHERS' ESTIMATIONS OF THEIR CHILDREN'S READING ACHIEVEMENT

Washington, Elois DeFrantz
THE CLASSIFICATION OF READING COMPREHENSION TEST ITEMS AND ITS RELATION TO THE PERFORMANCE OF SELECTED RACIAL GROUPS

White, Patricia Anne Davis
A CORRELATIONAL STUDY OF READING PERFORMANCE BETWEEN THE WIDE RANGE ACHIEVEMENT TEST AND TWO FORMS OF THE READING-SPELLING-VOCABULARY PROGRAM PLACEMENT TEST

A STUDY OF THE CLOZE PROCEDURE WITH NATIVE AND NON-NATIVE SPEAKERS OF ENGLISH Order No. 8005232

ALDERSON, JOHN CHARLES, PH.D. *University of Edinburgh (Great Britain)*. 1978. 341pp.

This study examined various aspects of the methodology of the cloze procedure to determine their effect on the nature of cloze tests. It was hypothesised that changes in the frequency of word deletion, in the difficulty of the original text and in the procedure used in judge acceptable restorations of the deleted word would produce significantly different cloze tests and would result in varying correlations with measures of English proficiency.

Three texts were selected and each was subjected to the deletion of every sixth, eighth, tenth and twelfth word, to give twelve cloze tests. Five procedures were developed to score the responses to these tests for the degree of similarity they showed to the deleted words.

The tests were administered to 360 adolescent native speakers of English and 360 adult non-native speakers of English who were pursuing further studies in Britain.

It was found that significant differences existed among cloze tests when deletion frequency was changed, but that some scoring procedures reduced this effect. The change in deletion frequency had no effect on the measurement of text difficulty, but significant interactions were observed among the three experimental variables. Different cloze tests gave unpredictably different measures of English proficiency. A study of identical deletions showed that no increase in the predictability of deleted words was gained by increasing context from five words to eleven words.

Since the quantity of context had no effect on predictability, it was suggested that cloze is essentially sentence-bound. The nature of the correlation of cloze with measures of English proficiency and the results of factor analyses suggested that cloze is a better test of syntax and lexis than of higher-order reading abilities. Implications for future use of the cloze procedure are presented and suggestions made for further research.

THE DIFFERENTIAL EFFECTS OF SYNONYMIC AND EXACT-WORD SCORING OF CLOZE TESTS UPON THE IDENTIFICATION OF INDEPENDENT, INSTRUCTIONAL, AND FRUSTRATION READING LEVELS Order No. 8001084

ARREOLA, Mona Jean, Ph.D. The Florida State University, 1979. 186pp. Major Professor: C. Glennon Rowell

The Problem

The early use of cloze testing has been recommended as a method for giving classroom teachers quick access to the information necessary for judging the ability of individual students to handle specific materials. Particularly in the content areas, such testing could help eliminate the extended observational period necessary for teachers to make informed judgments about students' instructional needs. However, restricting the scoring of these cloze tests to exact-word replacements, without allowance for synonymous answers, raises the possibility of underestimating students' actual reading performance levels.

Purpose

The purpose of this study was to investigate the degree to which synonymic and exact-word scoring of cloze tests constructed from a specific content text produced reading level placements – independent, instructional, frustration – that were in agreement with teacher judgment of the students' reading level placement in the text.

Procedure

Data were collected in December, 1978, from 67 sixth-graders in three intact classrooms. The three sixth-grade teachers were asked to rate their students' reading levels as independent, instructional, or frustration in the required social studies textbook. Using the Fry Readability Scale, a sixth-grade level passage of 250-300 words was selected from each fourth of the text. An every-fifth word, 50-item cloze test was constructed from each of the four passages. These were administered by the teachers, at the same time in the morning on four consecutive school days. Testing was done in groups of 22-26 students, with the three classes being tested simultaneously.

Each cloze test was scored twice by the investigator:

1. The number of exact-word replacements was counted and given two points each. The mean score from each subject's four cloze tests was translated into the corresponding comprehension level: independent, 60 – 100%; instructional, 40 – 59%; frustration, below 40%.

2. The non-exact responses were examined for synonymic acceptability. All acceptable synonyms were also given two points, which were added to the exact-word raw score. The mean synonymic score for each subject was then translated into the corresponding comprehension level.

Triserial r and Pearson Product-Moment correlations were used to determine the relationships between the exact-word cloze scores, the synonymic cloze scores, and teacher judgment. The McNemar Test for Related Samples, along with contingency tables, was used to examine the extent of agreement between the two scoring procedures and teacher judgment. All null hypotheses were tested at the .001 significance level.

Findings

Results showed that the use of synonymic scoring did produce reading level placements which were in closer agreement with teacher judgment than those determined by exact-word scoring. The McNemar Test yielded a highly significant difference between the two scoring procedures. An examination of this difference showed 49% more agreement with teacher judgment in identifying individual reading performance levels when synonymic scoring was used.

Correlational tests showed a moderate to high positive relationship between the variables: exact-word scoring and teacher judgment (r = .868), synonymic scoring and teacher judgment (r = .672), and exact-word scoring and synonymic scoring (r = .969).

Conclusions

Since the data showed that synonymic scoring of cloze tests produced significantly greater agreement with teacher judgment than exact-word scoring in identifying independent, instructional, and frustration reading levels in a specific content textbook, it is recommended that synonymic scoring be adopted as a routine procedure to be followed by classroom teachers when using cloze testing at the beginning of the school year to determine the reading performance level of individual students in the required content materials.

MONITORING AND FACILITATION OF READING GROWTH Order No. 8006577

AUSTIN, CECIL, PH.D. *University of Minnesota*, 1979. 63pp.

Purpose of the Study. This study was designed to assess the usefulness of one-minute samples of reading behavior taken from the local newspaper, as the dependent variable in the monitoring of reading growth. In addition, the study investigated the facilitative effect of the sampling method itself on reading performance. Statistical comparisons were made between mean reading rate performances and standardized reading test scores. Additional comparisons between pre and post measures were utilized to assess the facilitative value of the rate sampling method. Descriptive comparisons for individual pupils were also included.

Procedures. Subjects were forty-eight fourth and fifth graders from a school whose population was representative of a large urban school district. Their reading abilities were representative of the overall fourth and fifth grade population. Subjects were matched by grade, sex, and reading level, yielding a matched-pair design with one subject from each pair randomly assigned to an experimental group. All were administered a standardized reading test, pre and post. In addition, pre and post oral reading rates were taken. Between pre and post testing, the experimental group received additional reading rate sessions, the experimental treatment.

During a reading rate session, subjects were asked to select an article from the daily local newspaper, and to read orally for one minute. Then they were asked to evaluate their performance, which was intended to be facilitative, as it provided practice and feedback on the following activities: attention to task, distinctive feature focus, visual memory, auditory memory, and match-to-sample. This evaluation provided the following dependent variable measures: rate correct, rate missed, and rate of skips.

The data were analyzed in three ways. Pearson product-moment correlation coefficient were computed to assess the relationship of reading rate scores to a standardized test score. An analysis of covariance procedure adjusting for pretest performance was used to detect treatment effects on post measures. Finally, visual representation of data on the Standard Behavior Chart was used as a basis for interpreting the time series data.

Conclusions. The results of this investigation provide evidence that oral reading rates taken from the newspaper correlate significantly with standardized reading test scores. Somewhat weaker evidence was obtained that the method used in sampling reading performance facilitates reading growth.

Average reading rates, both pre and post were found to be positively correlated with total standardized test raw scores. The same relationship was found to exist between hit rates and all standardized sub-tests. Hit rates correlated more highly with the Passage Comprehension scores than they did with the Word Attack scores. Skips and misses tended to show low negative relationships with subtest and total test scores. The use of one minute samples of oral reading rate as a dependent variable in the monitoring of reading growth was supported by the data.

When groups were compared on adjusted post test measures, using standardized test raw scores, a treatment effect was obtained. This finding indicated that reading growth occurred which could be attributed to the reading rate sessions, and supports the contention that the sampling procedures improve readng. An analysis of reading rates failed to show a difference between groups, and therefore, did not lend any additional support to this finding.

In light of the widespread criticisms of contemporary monitoring and assessment procedures, these findings may have some implications for those involved in fostering the acquisition of reading skills. Should further research continue to varify these findings, many of the major criticisms of formal reading tests could be avoided.

THE EFFECT OF IMMEDIATE ITEM FEEDBACK ON THE RELIABILITY AND VALIDITY OF VERBAL ABILITY TEST SCORES

Order No. 8008637

BENSON, GORDON GUY EDWARD, PH.D. *The Florida State University*, 1979. 90pp. Major Professor: Dr. Howard Stoker

The purpose of this study was to investigate the effects of immediate item feedback (knowledge of results) on the reliability and validity of total test scores. Two types of feedback were studied: partial feedback (knowledge of correctness obtained by means of one attempt per item) and full feedback (knowledge of the correct response obtained by means of one attempt per item). Total feedback, or knowledge of the correct response obtained by answering until correct, was not involved.

Much of the previously published research on immediate item feedback appeared to be in need of larger sample sizes, and many designs did not appear to be capable of isolating the effects of feedback on mean test scores and reliability and validity coefficients of the test administered under feedback conditions. Their results were possibly confounded by using different response devices, time limits, numbers of attempts per item, and scoring strategies in the treatment and control groups.

Nine junior high schools in a large urban-suburban school district in the southeastern United States were selected using a stratified, random sampling procedure. Ninth grade students were assigned to cells in a 3 x 3, treatment-by-ability design and were tested on an adapted version of the SCAT-3B Verbal using Trainer-Tester response devices. Total scores of 2,023 students were analyzed with a non-orthogonal, ANOVA procedure and Scheffé comparisons. KR-20 reliability coefficients were analyzed using a *k*-sample test developed by Hakstian and Whalen (1976), and validity correlations with a subsequent reading achievement measure were analyzed with the usual tests for Pearson correlations.

Statistically significant main effects were found for treatment and ability, and the interaction was also significant. Examination for simple main effects indicated consistently lower means for the nonfeedback groups across ability levels and except for a reversal within the low ability level, full feedback means were generally lower than those for partial feedback. Reliability coefficients among the three treatment groups were statistically significant (partial feedback was greater than no feedback which was greater than full feedback), while the validity coefficients for partial and no feedback were significantly greater than that obtained for full feedback.

While a wealth of statistically significant findings were obtained, many of these significant differences were small. Criteria for judging educational, or practical, significance were discussed in terms of effect sizes (Cohen, 1969) and increased test length. After adopting suggested criteria, only one finding was judged to be educationally significant: for low ability students, there was a substantial increment in mean verbal ability scores in favor of full feedback over no feedback. Otherwise this study failed to show any substantial benefit or harm in students receiving knowledge of results while taking tests similar to those used in the study. The relevance of the study to previous research and suggestions for further research were also discussed.

PERFORMANCE VARIABLES IN ESSAY TESTING

Order No. 8007911

BIOLA, HEATHER ROBERTS, PH.D. *Georgia State University - College of Education*, 1979. 180pp.

Purpose. The purpose of this research was to determine what testing assignment conditions would allow students a fair opportunity to demonstrate their writing competence on an essay examination. The research was conducted in three phases: the first phase was an experimental investigation of the effect of topic assignments and time limits on essay test performance; the second phase was designed to find the best predictors of essay test scores; and the third phase was an analysis of the influence of testing conditions on the writing process and written products of four case study subjects.

Methods and Procedures. Three research methods were employed in this study to examine the different facets of performance reliability in essay testing. In Phase I a two-way factorial analysis of variance was used for the topic-by-time experiment. Forty-eight students wrote on a descriptive topic and another forty-eight students wrote on an argumentative topic. Half of the subjects in each topic group wrote the test essay within a time limit of 45 minutes, and the other half wrote within a time limit of 120 minutes. The dependent variable in this experiment was the experimental essay score holistically rated by three judges.

In Phase II a multiple regression equation was used to determine the best predictors of the experimental essay scores. The factors used as predictors included scores on another essay assignments, sex, race, parents' occupation, writing apprehension and reading habits. These factors have been studied in relation to writing performance in previous research.

Phase III included four case studies of individuals who participated in the first two phases of the project. The writing behavior of these subjects was observed during the testing period, and they were interviewed later. Finally, their oral language and their written diction, syntax and larger discourse were analyzed.

Results and Conclusions. The topic-by-time analysis of variance indicated that time was a significant factor: students who worked within the 120-minute time limit performed better than those who worked within the 45-minute time limit. However, topic had no significant effect on essay scores, and there was no observable interaction between topic and time.

The best predictors of the essay scores were scores on a similar essay test and the time limits assigned for the experimental essay. The full equation including all nine variables accounted for less than a third of the total variance of the experimental essay scores. Thus, this study found that essay test performance is unstable and hard to predict.

Finally, the writing process and written products of four students were described in detail. The two successful writers were standard dialect speakers who had the habit of criticizing and correcting their own writing during the writing process. Like one of Emig's case study subjects, these students avoided outlining their essays and preferred to depend on their own sense of cohesive organization. They believed they developed this sense of cohesion from their reading experiences. The unsuccessful examinees were nonstandard dialect speakers, who were unsure about written usage. They made outlines for their essays but failed to use the transitional and referencing devices which are necessary for a cohesive written text. All four case study subjects seemed to perform better when they had a longer time period to pursue the writing process.

A PROCESS MODEL FOR IDENTIFYING AND SELECTING PROFICIENCY STANDARDS IN THE BASIC SKILLS OF READING, WRITING, AND MATHEMATICS Order No. 8009734

BUCHER, LOUIS DAVID, ED.D. *University of San Francisco*, 1979. 250pp. Chairperson: Laurence A. Bishop

The purpose of this study was to identify and list in order of priority a set of competency goal statements in the basic skills of reading, writing and mathematics through a process of parent, student, staff and community involvement in order to comply with California's pupil proficiency laws as specified in Sections 51215 through 51225 of the *Education Code* of the state of California. Specifically, the study sought to identify and define competency goal statements through a democratic process of shared decision making, and finally the study endeavored to generate a reliable data base for use in the construction and validation of proficiency test items.

Initially, a continuum of basic skills and abilities in the areas of reading, writing, and mathematics was prepared by a group of students, parents, faculty members and community representatives. This input was screened, edited and translated into a questionnaire containing a series of words and phrases representative of the various functional applications of reading, writing, and mathematics. For each skill area, a five point Likert scale ranging from a high of 1 (critical for getting along in society) to a low of 5 (unnecessary for getting along in society) was used. The questionnaire was administered to a representative sample of parents, students, faculty members and community representatives. The ratings of the respondents were intercorrelated and factor analyzed in order to ascertain the factor structures (clusters) for reading, writing, and mathematics from which competency statements were inferred. The mean ratings of the total population sample and of each division in the sample were computed to determine the overall priorities and the priorities within each subgroup. Finally, using the domain cluster means of each skill area as dependent variables and variations of the subgroup populations as independent variables, group differences (variance) in perceived priorities were analyzed.

A total of fourteen competencies (five each in reading and mathematics and four in writing) were generated from this study for use in the Eureka High School District. Within each skill area, competencies range from basic or concrete skills and abilities to more complex or abstract skills and abilities. This general pattern of competency organization is consistent across skill areas and represents a natural hierarchy from skill orientation and acquisition to skill building and skill application. As perceived by the total community, these competencies represent the knowledge, skills and abilities that students should acquire before graduating from high school.

Differences in priority ratings between and within the various subdivisions of the study sample (parents, pupils, faculty, and community representatives) did occur, but the ordering of priorities was the same across subgroups. In relation to the priorities assigned by each subgroup, parents consistently assigned the highest degree of importance to the competencies within each skill area, while faculty members generally assigned the lowest degree of importance. Additionally, there were generally no significant differences in the priority ratings assigned by parents and community representatives of different occupational classifications to the competencies within each skill area. In other words, parents and community members, regardless of socio-economic status, assigned the same relative degree of importance to the knowledge, skills, and abilities contained in each competency.

The findings of this study have resulted in the following implications: (1) group decision making can be democratic; (2) competencies represent skills which are related by function; (3) competency domains provide a guide for curriculum development; (4) competency domains provide a guide for student evaluation; (5) schools should ensure student competency in the basic skills before graduation; and (6) occupational groups are in agreement on educational priorities.

A STUDY OF THE CLOZE PROCEDURE UNDER THREE DIFFERENT CONDITIONS AT TWO COGNITIVE LEVELS Order No. 8010575

COTHRAN, ANN HOLLINGSWORTH, PH.D. *University of Georgia*, 1979. 113pp. Director: Ira E. Aaron

The purpose of this study was to investigate the relationship between scores on the *Group Embedded Figures Test (GEFT)* and cloze tests based on structural, lexical, and every fifth word deletion patterns. Two hundred seventy-two fifth graders reading at the middle of the fourth grade level or above were given the *GEFT*, then divided into field independent groups on the basis of their scores. Twenty-four field independent boys were assigned to either structural, lexical, or every fifth word deletion treatment, and twenty-four field independent girls were randomly assigned to these three treatment categories also. Forty-eight field dependent boys and girls (24 boys and 24 girls) were also randomly assigned to the three treatment categories. All subjects were given three cloze passages to complete. Data were analyzed by two ANCOVAs, with IQ the covariate. Results indicated that field dependence was not significantly related to cloze performance ($F = .21$, $p < .05$). Type of deletion pattern was a significant factor ($F = 131.5$, $p < .05$), and pairwise multiple comparisons indicated that the most difficult cloze to complete was the lexical, with the every-fifth word significantly easier and the structural the easiest to complete.

The findings of this study suggest that *GEFT* performance is not significantly related to cloze test performance when intelligence is controlled. Results confirm previous findings of differential difficulty levels of the cloze when deleted by various patterns. As in past studies, the order of difficulty was lexical cloze most difficult, then every fifth word, and structural cloze the easiest.

SCHEMA THEORY AS A THEORETICAL FRAMEWORK FOR THE CLOZE PROCEDURE Order No. 8005030

DAVENPORT, BETTY MUNCH, PH.D. *The University of North Carolina at Chapel Hill*, 1979. 168pp. Supervisor: Dixie Lee Spiegel

The effects on cloze test performance of reader proficiency and sex, familiarity of topic, explicitness of writing style, and type of deletion (content or function word) were investigated in order to gain insight into the reading comprehension process. Schema theory was used as the organizing framework for the study. Two levels of each of the three prose variables were used for devising the four experimental cloze tests. The tests were written at a ninth-grade readability level. The subjects were 180 students of four reading proficiency levels from three North Carolina community colleges.

The high proficiency readers achieved the highest total cloze test scores and the very low proficiency readers achieved the lowest scores. For the group as a whole, twice as many function words as content words were given verbatim, but the total test scores of the high proficiency readers were comprised of a larger proportion of content words than were the total scores of the less proficient readers. The high proficiency readers were not significantly influenced by topic or writing style, but the less proficient readers were. The least proficient readers were the only ones whose cloze test performance was favorably influenced more by the familiarity of the topic than by the writing style. The sex of the readers did not influence the test scores significantly.

The duplication of certain deletions and the repetition of deleted words in the intact portion of the test were identified as confounding variables operating within the tests.

The results of the project suggested that further investigation of schema theory as an organizing framework for the study and interpretation of the cloze procedure (when used for devising reading tests and for measuring prose readability) is probably warranted.

THE VERIFICATION OF THE TEST OF AFFIXES IN SYNTACTIC STRUCTURES: A STUDY OF DERIVATIONAL MORPHOLOGY AS A LANGUAGE CORRELATE FOR COLLEGE-LEVEL READING PROFICIENCY Order No. 8000782

DOGGER, Barbara Teresa, Ph.D. North Texas State University, 1979. 253pp.

A lack of research in adult literacy for both native speakers and speakers of English as a second language led to the development of the Test of Affixes in Syntactic Structures (TASS) for use in a pilot study (Dogger, January 1978) in which knowledge of derivational morphology was tested to determine its possible relationship with reading for English as a second language students. The test was developed according to the following theoretical constructs.

(1) Knowledge of language as demonstrated by the receptive language skills of listening and reading is tested efficiently by focusing upon a specific language correlate. (2) The language correlate of derivational morphology was selected for TASS for several reasons. Because morphology is the focal point of all language systems--syntax, semantics, and phonology (Chomsky, 1958, 1965, 1970; Halle, 1973; Jackendoff, 1975; and Aronoff, 1976)--a person's knowledge of these systems is most efficiently determined through a test of morphology. The orthography of English optimally reflects its morphological nature (N. Chomsky

and Halle, M., 1968; C. Chomsky, 1970) in these derivational forms. (3) Knowledge of spoken and written morphology is tested in TASS. Part I tests a person's knowledge of the complex phonological system of English and mastery of English orthography as it represents both the graphophonic and the morphophonemic level of language. Part II tests this knowledge of English orthography as it specifically signals derivational affixes and a person's knowledge of the syntactic and semantic function of nonsense words as signaled by these affixes.

Test construction was followed by a thorough verification procedure which is the purpose of this study. In September 1978 the following measures of test strength were established: construct validity, content validity, item difficulty, item discrimination, internal consistency, rational equivalence, and concurrent validity. The degree of relationship between reading proficiency, as demonstrated by subject performance on the Iowa Silent Reading Test, Level III (ISRT, III), and knowledge of derivational affixes, as demonstrated by subject performance on TASS, was also established.

Results show that successful performance on the ISRT, III includes reading strategies beyond those required for successful performance on TASS. In other words, mastery of language structures as represented by English orthography is necessary but not sufficient for college-level reading proficiency.

Recommendations for further research include the following. (1) TASS should be administered along with a test of general comprehension to subjects identified as beginning-to-intermediate readers in order to determine the point at which mastery of derivational forms is most crucial to reading proficiency. (2) Because the two subtests included in TASS measure specific language skills, each subtest should be investigated to determine its usefulness as a language subskill test in diagnosing strengths and weaknesses when reading difficulty is apparent. Testing should take place with subjects reading at various grade levels in order to determine the level at which its usefulness is optimal.

A COMPARISON OF TECHNIQUES FOR THE IDENTIFICATION OF CHILDREN WITH READING DISABILITIES

Order No. 8005036

DOMINICK FIELDS, NANCY ANNE, PH.D. *The University of North Carolina at Chapel Hill*, 1979. 57pp. Supervisor: Walter B. Pryzwansky

This study compared the use of three identification formuli and two levels of assessment instruments on a single school sample to determine first, whether the same children were identified as reading disabled and secondly, how the groups of children identified as reading disabled compared on a number of demographic and behavioral variables.

The third grade (121 children) of a small North Carolina school system served as subjects. The Slosson Intelligence Test (SIT), Slosson Oral Reading Test (SORT), Otis-Lennon Mental Ability Test, and Iowa Test of Basic Skills (ITBS) were administered. Teachers completed the Devereux Elementary School Behavior Rating Scale. Six groups of 12 children were identified. Three groups were identified using individual test data and the following methods: (1) Years Below Grade Level, (2) Bond and Tinker Expectancy Formula, and (3) Z-score Discrepancy Formula. Three groups were identified using group test data and the three methods described above.

Comparisons of the groups were made on the following criteria: Wechsler Intelligence Scale for Children (WISC) Full Scale IQ, WISC Verbal IQ, WISC Performance IQ, age, years in school, educational level of parents, sex, race, and the Devereux Elementary School Behavior Rating Scale characteristics. Using individual test data the Years Below technique and the Bond and Tinker formula identified children with lower IQs than did the Z-score formula. Using group test data the Years Below method showed significantly lower IQs than did the Bond and Tinker or Z-score methods. Group tests identified significantly higher numbers of females than individual tests for all three methods. None of the remaining demographic characteristics showed significant discrepancies. There were minimal discrepancies in the behavioral characteristics among the six groups.

Individual and group tests did not identify the same children. For each of the identification formuli there was negligible overlap between those identified by individual assessment data and those identified by group assessment data. The results of comparisons both across and within methods would suggest that there is reason to question the validity of using group test data to identify reading disabled children. Not only do they not identify the same children, but they also do not identify groups with comparable IQ and reading test means.

Findings in this research indicate that the levels of assessment instruments are not interchangeable and that the identification methods used are not interchangeable.

THE DEVELOPMENT OF A DIAGNOSTIC READING ATTITUDE ASSESSMENT FOR SECOND GRADE

Order No. 8000993

DRYDEN, Beverly Joan, Ed.D. University of Georgia, 1979. 221pp. Major Professor: Ira E. Aaron

The purpose of this study was to develop a reliable and valid diagnostic reading attitude assessment which measures the instructional and recreational reading attitudes of second grade children within a school setting.

Content validation procedures were implemented in which photographs and written questions representing instructional and recreational reading situations were judged by teachers of second grade children. Forty of these photographs and questions were rated by judges as appropriate. The makeup of the Assessment included these 40 items. The items were divided into eight subtests: (1) Reading Aloud, (2) Working in a Reading Group, (3) Reading Seatwork, (4) Taking Reading Tests, (5) Working in a Reading Workbook, (6) Instructions Given by the Teacher, (7) Recreational Reading Within the Classroom, and (8) Recreational Reading in the Library.

Three formats were then developed in order to determine the format that was most reliable and valid. Format PQ of the Diagnostic Reading Attitude Assessment contained black and white photographs and written questions representing instructional and recreational reading settings. Format P contained only the black and white photographs. The written questions were deleted from this format. Format Q contained only the written questions. The photographs were deleted. For each format, the same 40 questions were read orally to the students. Students marked each item on the same three response scale.

The subjects in the preliminary tryout of the three formats and the 40 items consisted of 141 second grade students.

Item-total test correlation coefficients revealed that the Question Format was the most valid to use with second grade children, yielding more items with reasonably good to very good item-total test validity coefficients than did either the PQ or P Format.

A comparison of the three subtest-total test correlation coefficients for each subtest in three formats revealed that the values of the coefficients were higher in the Q Format subtests than in Formats PQ and P. A comparison of the three reliability coefficients in each format revealed that the value of the coefficient was higher in the Q Format (.9414).

From the results of the data, the Question Format was selected as the most appropriate format to use in the final research.

After the preliminary data were analyzed and the Question Format chosen, a total of 37 items was included in the revised Diagnostic Reading Attitude Assessment. These items represented instructional and recreational reading situations within a school setting and were divided into the same eight subtests. Children in the final research project consisted of 135 second grade students. Item-total test correlation coefficients revealed that 89 percent of the items had reasonably good to very good item-total test validity coefficients. The values of the subtest-total test correlation coefficients were well within the acceptable validation range. An estimate of internal consistency (.8973) of the total test indicated that the revised Assessment was reliable.

Criterion validation analysis of the relationship between scores of students and scores of teachers indicated a weak positive relationship in the subtests, Reading Instruction and Classroom Recreational Reading. A negative relationship occurred in the subtest, Reading Group. The remaining subtests had no relationship since they had correlation coefficients of near zero.

The conclusions reached included: (1) The items and subtests within the Assessment seemed to provide the teacher with diagnostic information concerning a child's attitude to-

ward six areas of instructional reading and two areas of recreational reading within a school setting. (2) The Question Format seemed to be the format that was most valid and reliable. The format and response options seemed easy for children to mark and the Assessment seemed easy to score. (3) The profile sheet on the cover seemed to provide the examiner with a quick analysis of a child's attitude toward eight areas of instructional and recreational reading.

A COMPARISON OF QUANTITATIVE AND QUALITATIVE METHODS OF ANALYZING THE ORAL READING BEHAVIOR OF FOURTH-GRADE STUDENTS

Order No. 8012025

FEDORKO, FREDERICK JOHN, PH.D. *The University of Akron*, 1980. 238pp.

The purpose of this study was to estimate the efficacy of a qualitative analysis of oral reading behavior as compared to the traditional quantitative analysis of the Informal Reading Inventory in predicting the reading effectiveness of fourth-grade students and in determining their appropriate instructional level. In both instances, the criterion used was teacher judgment. The study also attempted to establish known group validity for both the quantitative and qualitative procedures. Three broad questions were posed in this study. (1) Will the quantitative procedures of the Informal Reading Inventory or the qualitative procedures of Miscue Analysis more accurately predict the reading effectiveness of fourth-grade students as determined by teacher judgment? (2) Will the quantitative procedures of the Informal Reading Inventory or the qualitative procedures of Miscue Analysis more accurately predict the appropriate instructional level of fourth-grade children as determined by teacher judgment? (3) Will either of the above procedures be more effective in discriminating good, average, and poor fourth-grade readers as determined by an overall rating of reading effectiveness given to them by their teachers? The graded passages of the Analytical Reading Inventory (Woods & Moe, 1977) were read by the entire population of 99 fourth-grade students, and an informal questionnaire developed by the investigator was used to determine teacher judgment of reading effectiveness and instructional level. When testing was completed the oral reading of each student was analyzed both quantitatively and qualitatively using the procedures of the *Informal Reading Inventory* and the Reading Miscue Inventory (Y. Goodman & Burke, 1972) respectively. When the scoring and analysis of each subject's reading were completed, estimates of inter-judge reliability were determined. The research design of this study was *ex post facto* research with hypotheses and tests for rival hypotheses. The statistical procedure chosen to analyze the data was Multiple Linear Regression. From the data gathered in this investigation, 25 hypotheses and 36 alternative hypotheses were formulated and tested at the .05 level of significance for a two-tailed test. The results supported the ability of both quantitative and qualitative procedures to predict teacher judgment of reading effectiveness and instructional level. The results did not find either the quantitative or the qualitative procedures to be effective in discriminating good, average, and poor readers. The estimates of known group validity for both procedures were low. The inter-judge reliability estimates for both the quantitative and qualitative procedures were high when five judges were used. The only exception was the retelling score of the qualitative analysis. It was suggested that the low reliability for the retelling score may be due to its uncommon usage. An analysis of the data generated the following broad conclusions: (1) both quantitative and qualitative procedures were good predictors of teacher judgment of reading effectiveness and instructional level; however, the qualitative procedures are slightly superior to the traditional quantitative procedures; (2) the qualitative procedures generate more specific and detailed information upon which pedagogical decisions might be based; (3) neither the quantitative nor the qualitative procedures have a significant measure of known group validity; (4) the traditional analysis of the Informal Reading Inventory adequately represents a student's performance in the classroom as determined by teacher judgment; (5) the poor readers in this study were weak in phonics and did not rely heavily upon grapho-phonic cues; (6) poor readers were unable to use syntactic and semantic information effectively; (7) average readers used grapho-phonic information, but not to the exclusion of meaning; and (8) good readers did not rely heavily upon grapho-phonic information, and were proficient in using syntactic and semantic cues.

THE CONSTRUCTION OF AN INSTRUMENT TO MEASURE ELEMENTARY TEACHERS' KNOWLEDGE OF A SUBSKILLS APPROACH TO TEACHING READING

Order No. 7928645

HARPER, BETTY JEAN, PH.D. *The University of Wisconsin - Madison*, 1979. 190pp. Supervisor: Professor Wayne Otto

This is a report of the procedures followed in constructing an instrument to assess elementary teachers' knowledge of a subskills approach to teaching reading. In the first phase of instrument development an outline of test content was developed, and was submitted to four reading authorities for evaluation. The second phase involved writing and selecting items for the preliminary test form. Based on the test outline, a pool of multiple choice items was generated and subsequently reviewed by persons knowledgeable in the subject matter area. The items were then revised and/or replaced, based on the reviewers' comments. A further check on content validity was made when knowledgeable persons keyed the items and furnished information for improving the preliminary test form. The third phase of the test construction process involved: (1) field testing the preliminary test form with elementary teachers, (2) analyzing and revising the items based on the field tryout data, and (3) reassessing content validity of the revised instrument. In the final phase of test construction the revised instrument was subjected to a field tryout with 104 elementary teachers. The resulting data were examined to determine reliability and content validity of the present instrument. In addition, a one-tail t-test was used to determine whether differences between the group mean scores for teachers with zero or one reading course and teachers with three or more reading courses were significant at the .05 level. The resulting separate variance estimate indicated a t-value of -.31, less than the critical value of t (1.98) required for significance at the .05 level. The correlations between number of reading courses and years of teaching experience to test scores were also examined. The Pearson product-moment correlation between number of reading courses and performance on the instrument was computed at .14 while the correlation between years of teaching experience and performance on the instrument was -.31. An examination of data from the second field tryout indicated that the present instrument, with a Hoyt internal consistency reliability coefficient of .82 and a standard error of measurement of 3.29, may be considered adequate for interpreting individual measurement. Further, a review of the test development process indicated that it is highly probable that the present instrument demonstrates content validity. This conclusion is based on the cumulative evidence that content validity was built into the instrument from the beginning of the test development process by systematic, logical analysis for determining the adequacy of the items for measuring content in the test outline.

PREDICTING SUCCESS ON A MINIMUM COMPETENCY EXAMINATION

Order No. 8011413

HENDERSON, DONALD JOSEPH, ED.D. *Loyola University of Chicago*, 1979. 136pp.

The purpose of this study was to investigate the relationship between four variables and performance of 10th graders on a minimum competency examination in reading. The Reading Proficiency Examination is administered by the Gary, Indiana public schools as one requirement for high school graduation.

This study was designed to ascertain whether these variables can be used as an early warning system to identify students who might fail the competency examination and therefore benefit from remediation at the middle school level. To this end the following null hypotheses were generated: (1) School attendance does not significantly predict performance on the Reading Proficiency Examination. (2) Iowa Tests of Basic Skills reading comprehension subtest scores do not significantly predict performance on the Reading Proficiency Examination. (3) Sex of the student does not significantly predict performance on the Reading Proficiency Examination. (4) Socio-economic status does not significantly predict performance on the Reading Proficiency Examination.

The subjects used in this study were 203 students who were tenth graders during the 1977-78 school year from two high schools. Multiple regression analysis was the major tool of statistical analysis.

The results of this analysis lead to the rejection of Hypotheses 1, 2, and 4 and the acceptance of Hypothesis 3. Moreover, the analysis revealed a significant level of intercorrelation between 3 measures of basic skills: reading, mathematics and written communication. Finally, the analysis indicates that the Iowa Test of Basic Skills secondary reading percentile for students in middle school significantly predicts performance on the three aforementioned measures of basic skills, and therefore, can be used as the basis of an early warning system to detect students who need help with basic skills prior to entry into high school.

A COMPREHENSION STUDY OF THREE VERSIONS OF THE GOSPEL OF MARK AND THEIR APPROPRIATENESS FOR USE BY SIXTH GRADERS WITH VARIOUS READING ABILITIES

Order No. 8006272

HERZOG, KATHARINE DORA, PH.D. *The Florida State University*, 1979. 97pp. Major Professor: Edwin H. Smith, Ph.D.

The purpose of the study was to assess the comprehensibility of three versions of the Gospel of Mark and the appropriateness of using them with sixth grade students. The subjects involved in the study were two hundred and seventy-nine sixth grade students from six private Christian schools in Florida.

The three versions of the Gospel of Mark used in the study were the King James version, the Good News version and the New International version. The readability level of each version was measured by the Flesch Reading Ease Formula, the Dale-Chall Readability Formula and the Smith Adaptation of the FORCAST Readability Formula.

Reading ability levels of the subjects were determined by reading comprehension scores from the Stanford Achievement Test.

A thirty-two question multiple-choice comprehension test was developed by the examiner. The test measured reading comprehension of the specific passages of Mark used in the study. It was administered to the sixth grade subjects. The reliability of the test, as measured by the split-halves method and the Spearman-Brown formula, was found to be .83.

Findings of the study were displayed in a series of tables. A comparison of the tables was made. Results of the readability formulas indicated the level of the King James version of Mark to be Grade five or six; however, 0 percent of the students at those reading levels were able to answer 75 percent of the comprehension questions related to the measured passages. Results of the readability formulas indicated the level of the New International version of Mark to be Grade five or six; however, only 4 percent of the students at those reading levels were able to satisfactorily answer 75 percent of the comprehension questions related to the measured passages. The results of the readability formulas indicated the level of the Good News version of Mark to be Grade five or six; however, only 19 percent of the students at those reading levels were able to satisfactorily answer 75 percent of the comprehension questions related to the measured passages. The lowest reading level at which 50 percent of the students could satisfactorily read and comprehend the Good News version of Mark appeared to be the ninth-tenth grade level; for the King James and New International versions of Mark, it appeared to be above tenth grade.

The investigator concluded that none of the three versions of the Gospel of Mark used in the study was appropriate for use with sixth grade students reading on grade level. She also concluded that the readability formulas used were not valid for use with this type of study. The findings of her investigation implied that reading passages should not be assigned on the basis of results of readability formulas alone.

That intermediate grade readers are unable to adequately comprehend some of the more popular versions of Scripture was clearly supported by the study. This would indicate a strong need for a rewritten version of the Bible in a language more appropriate for children.

A COMPARISON OF THE READING SUBTEST OF THE CALIFORNIA ACHIEVEMENT TESTS, THE GRAY ORAL READING TEST, AND THE PUBLISHERS' PLACEMENT TESTS FOR THE STATE-ADOPTED BASAL READING SERIES

Order No. 8001295

HODGES, Jane Clayton, Ed.D. The University of Mississippi, 1979. 99pp. Director: Dr. Minnie Y. Peaster

The purpose of this study was to obtain evidence concerning the extent to which the level of reading achievement as measured by standardized reading achievement tests and oral reading tests differed from the levels determined by basal placement tests accompanying the basal reading series. More specifically it was designed to test two null hypotheses which stated that there would be no significant differences between the levels of reading achievement as measured by the California Achievement Tests or the Gray Oral Reading Test and the placement level as determined by the placement test for the basal reading series being used.

A total sample of 240 subjects was drawn from ESEA Title I remedial reading classes at four Mississippi public schools. Thirty subjects from the third and the fifth grade classes were randomly selected.

The reading subtest of the California Achievement Tests, 1970 Edition, Form A, Levels 2 or 3; the Gray Oral Reading Test, Form A; and the placement test accompanying the basal series being used at the school were administered to each subject. All scores were converted to grade equivalent scores.

All data were processed mechanically and analyzed at the .05 level of significance. Both null hypotheses were rejected, based on t ratios calculated for all selected test pairs, which indicated that 11 of the 16 pairs were significantly different. Coefficients of correlation (.0335 to .7598) indicated very low correlation. A comparison of means showed significant differences between the means of the California test (.1500 to 2.3400) and the Gray test (.2250 – 1.0800) when compared to the means of the basal tests. Results of a comparison of individual scores indicated that only a small percentage of the subjects were placed at the same basal level by either the California test (16%) or the Gray test (20%) as compared to basal test placement. The California test placed 42% of the subjects within one half year higher or lower than the basal test, including those with the same placement. The Gray test placed 50% withon one half year higher or lower.

The following conclusions were drawn in accordance with the purposes of this study:

1. The reading subtest of the California Achievement Test, 1970 Edition, Form A. Levels 2 or 3, is probably not an appropriate test instrument for placing elementary remedial reading students at their instructional reading levels in the basal reading series studied.
2. Standardized reading achievement tests should be used with discretion for placing elementary remedial reading students at their instructional reading levels.
3. The Gray Oral Reading Test, Form A, is probably not an appropriate test instrument for placing elementary remedial reading students at their instructional reading levels in the basal reading series studied.

BEYOND READABILITY: MEASURING THE DIFFICULTY OF TECHNICAL WRITING

Order No. 8004761

HULL, LEON CLINTON, JR., PH.D. *Rensselaer Polytechnic Institute*, 1979. 193pp. Supervisor: David L. Carson

The application of readability formulas to technical writing poses two problems. First, most formulas rely on a generalized measure of word difficulty (number of letters or syllables per word) which makes no allowance for familiarity with a technical vocabulary. These formulas therefore tend to rate technical material as more difficult than it actually is for the intended audience. Second, since traditional formulas employ indices that predict but do not in themselves cause difficulty, they are inapplicable as a tool for improving technical writing.

To be useful to technical writers, a formula should be generally applicable to technical material; that is, it should employ variables that are relatively independent of vocabulary difficulty. A formula for technical writing should also use indices that stand in a causal relationship to difficulty, thus providing writers with an editorial tool for revising their material to make it more readable. The development of a formula for assessing and improving the readability of technical writing was the goal of this study.

The literature on readability indicated that the difficulty of a written message depends in large part on the number and type of syntactic structures used to modify the ideas conveyed in the message. Moreover, the degree of modification appeared to be independent of vocabulary load and a possible cause of difficulty. A syntactic analysis of criterion passages drawn from technical materials identified several indicators of modifier load that have a substantial correlation with textual difficulty.

An empirical study was therefore designed to test the hypothesis that an increase in these indicators of modifier load *causes* an increase in passage difficulty, as indicated by a decrease in reader understanding on a standard cloze test. A second purpose of the study was to verify the value of audience analysis in technical writing by assessing the effect of verbal skill, familiarity, and interest on reader understanding. The subjects were 107 undergraduate students at a major technological institute.

The principal hypothesis dealing with modifier load received some strong support. Test results showed that a 50-percent increase in the density (number) of prenominal modifiers resulted in a significant decrease in scores on the cloze test. Increased prepositional density had no significant effect on reader understanding, indicating that the prepositional phrase could be an acceptable alternative to prenominal modifiers in revising technical material.

The tests also supported the current emphasis on audience analysis. When averaged or aggregated across a set of passages, the three reader variables showed a significant correlation with reader understanding. However, these same variables were only marginally effective in predicting the success an individual reader would have with an individual test passage.

The two syntactic variables (other than vocabulary load) which correlated most highly with difficulty provided the elements for a subsequent regression analysis based on the criterion passages. The product of this analysis was a prototype formula applicable to technical writing.

The major conclusion to be drawn from this study is that it is feasible to develop a formula which will serve as a means for assessing the readability of technical writing and as a tool for revising technical material to make it more readable. In the initial version of the formula presented here, syntactic length (number of words per sentence) combines with modifier load (density of prenominal modifiers) to account for 47.87 percent of the total variation in passage difficulty. Although not quite as powerful as the recalculated Flesch formula recently endorsed by the Department of Defense for use in the preparation of technical manuals, this prototype formula is fully as accurate as most of the traditional readability formulas.

A STUDY OF THE DIFFERENCES BETWEEN READING SCORES: AS ACHIEVED ON SILENT INFORMAL READING INVENTORIES AND ORAL INFORMAL READING INVENTORIES BY ELEMENTARY STUDENTS

Order No. 8000072

JORGENSON, Jeffrey Lee, Ed.D. Brigham Young University, 1979. 80pp. Chairman: Floyd Sucher

The purpose of this study was to determine if a difference of instructional reading level scores was produced when comparing oral IRI results with silent IRI results. Twenty-four pupils in each of the second, fourth and sixth grade levels were randomly selected and administered four methods of inventories: a McCracken silently, a McCracken orally, a Silvaroli silently, and a Silvaroli orally.

Based upon the results of a statistical analysis, it was concluded that: (1) second grade students achieved higher instructional reading levels by taking an oral IRI; (2) the Silvaroli inventory resulted in higher instructional reading levels than the McCracken inventory; (3) sixth grade students have learned how to read better than fourth grade students, and both have learned how to read better than second grade students; and (4) second grade female students are better readers than male second grade students.

TWO MODES OF ANALYZING TEACHER AND STUDENT ATTITUDES TOWARD WRITING: THE EMIG ATTITUDE SCALE AND THE KING CONSTRUCT SCALE

Order No. 8000863

KING, Barbara Lynn, Ed.D. Rutgers University The State University of New Jersey (New Brunswick), 1979. 217pp. Chairperson: Robert P. Parker

The purpose of this study was twofold: to establish what attitudes toward writing were held by teachers and students, and to determine if these attitudes could be changed through teacher participation in the New Jersey Writing Project's 1977 Summer Writing Institute.

Eighteen experimental teachers, seven control teachers, 1093 experimental students and 601 control students completed one or more test measurement instruments. All subjects completed scales designed for this study to ascertain their attitudes toward writing. All teachers completed the Emig Teacher Attitude Scale. The Emig Student Attitude Scale was completed by 1093 experimental students in the fall; 952 completed the same form in the spring. The Emig Student Attitude Scale was completed by 601 control students in the fall and 449 control students responded to the same form in the spring. In addition, 263 control students completed the King Construct Scale.

Both the fifty-two item Emig Student Attitude Scale and the fifty item Emig Teacher Attitude Scale presented items in three categories: preferences for writing, perceptions of writing and processes of writing. For each item the respondent circled one of five responses on a Likert-type scale. The King Construct Scale, based on George Kelly's psychology of personal constructs, presented twenty-five bi-polar constructs separated by a five point scale. The twenty-five bi-polar constructs represented student preferences for: source, purpose, audience and mode.

The analysis of the results from the three test measurement instruments provided data which were analyzed in tabular fashion. Data were analyzed comparing students in the experimental group and students in the control group, fall and spring; teachers in the experimental group and teachers in the control group; and students, by sex.

The analysis of the Emig Student Attitude Scale showed that students in the experimental group exhibited a more positive attitude toward writing than students in the control schools. Experimental students, at the end of their English course, reported a greater preference for writing, reported writing more on their own, and reported a greater preference for what they wrote.

The female students in both the experimental and control groups reported a greater preference for writing and reported writing more on their own than the males in both the experimental and control groups. This sex difference was noted in both fall and spring.

The analysis of the results of the Emig Teacher Attitude Scale showed that experimental teachers reported writing more and viewed the teaching of writing differently than control teachers. Control teachers placed greater importance on correctness of spelling and grammar, while experimental teachers placed greater importance on content.

The analysis of the results from the King Construct Scale showed that students prefer self as the source of their writing, personal choice of topic, self as audience, and original writing as a mode. They also indicated a preference for graded as opposed to ungraded papers.

The study concludes that teacher participation in the 1977 Summer Writing Institute of the New Jersey Writing Project led to a positive change in attitudes toward writing of the participating teachers and their students.

TEACHERS' COMMENTS ON STUDENTS' WRITING: A CONCEPTUAL ANALYSIS AND EMPIRICAL STUDY

Order No. 8003942

KING, JEAN ANNE, PH.D. *Cornell University*, 1979. 158pp.

The problem addressed by this thesis was that of improving the nearly universal practice of teachers' commenting on students' writing, a traditionally negative practice. An analysis of the concept of composition annotation provided a framework for answering three questions: (1) What do we now know about the effect of comments on writing; (2) What evidence is there that rule explicitness affects student writing; and (3) What are the implications for research and practice?

The expansion of an empirical analysis by William McColly and a summary of research by Daniel Wolter and Walter Lamberg suggested eight dimensions of composition annotations: source; delivery mode; timing; purpose; form (appearance, length, placement, sentence type, and number); content (tone, voice, vocabulary, relation-to-task, level, and rule explicitness); textual referent (focus, correctness, quantity, and identification); and value (appropriateness and adequacy). These dimensions can be used to analyze sets of comments on papers as well as variables used in research.

Past research on the effect of comments, suffering from both conceptual and methodological problems, has resulted in few general claims. There are two exceptions to this; research has shown that positive comments lead to an improved attitude toward writing (although not necessarily to improved writing) and that more comments do not improve writing attitude or ability better than fewer comments. The studies which have included comment explicitness as a variable have suggested little about its effect on students.

The approach used in the empirical study differed from that of previous research in that it examined the process of student's reaction to comments. Comments differing in rule explicitness were placed on three paragraphs, each of which had been altered to contain the same nine errors. Seventy-two students were given two paragraphs, one marked and one unmarked, and asked first to write what the comments on the first paragraph meant to them

(i.e. *understanding*) and then to *find* and correct similar errors in the second paragraph (i.e. *fixing*). They were also asked to indicate which comments they felt would help them improve their writing.

Analyses of variance showed the following significant main effects: ability group on the dependent variables of understanding, finding, fixing, and helpfulness of comments; comment explicitness and error type on understanding comments; and error type on comment helpfulness. The lack of effect of the independent variables of comment explicitness on the finding and fixing variables was probably due to the shortness of the treatment and a lack of difference between two of the comment treatments for some errors. Analysis of the response patterns on the variables of understanding, finding, and fixing lent some support for a rule-based model of learning from comments; 80% of the respondents showed predicted patterns which were consistent with the order of understand-find-fix.

It was not surprising that better writers did all tasks better, nor that students misunderstood and found less helpful comments about certain errors. But the fact that students understood the actual correction of errors significantly less often than comments labeling errors or comments giving rules for correcting them suggests that mere correction may not be understood by students.

The conceptual analysis of composition annotation suggests variables for future research, while teachers can use the dimensions to determine the structure of their own comments. With refinements, the method used in this study may also be of value, leading to information both about the effect of comments and about how students learn to write as a result.

A COMPARATIVE STUDY OF SCREENING TESTS IN THE EARLY IDENTIFICATION OF DYSLEXIA

Order No. 8001014

KIRBY, Timothy Stephen, Ph.D. University of Georgia, 1979. 104pp. Major Professor: Warren C. Bonney

Thirty-one first grade pupils were randomly selected from a county school system in Georgia. Each pupil was tested using the Prereading Expectancy Screening Scale, the Meeting Street School Screening Test, Wechsler Intelligence Scale for Children, the Bender Gestalt, and the reading subtest of the Wide Range Achievement Test. The results of these instruments were correlated with each other and with the reading level as measured by the Scott Foresman Reading Systems which was used as the criterion reference. The results indicate that the screening instruments have high predictive validity of reading levels.

AN ANALYSIS OF THE ATTITUDES OF SENIOR HIGH SCHOOL STUDENTS TOWARD STUDENT-EVALUATED COMPOSITION

Order No. 8006700

KYHOS, RONALD MILTON, PH.D. *University of Maryland*, 1979. 129pp. Supervisor: Professor Leonard Woolf

Purpose. The focus of this study was to measure the differences among high school students in their general attitudes toward student-evaluated composition. The three areas making up the general attitude were defined as: (1) the involvement of the students in the process, (2) the additional time for teachers to help students, and (3) the better learning of writing skills within the process of student-evaluation of compostion. A Total Attitudinal Score and scores on these three areas were classified by: (1) sex, (2) scores on a standardized test dealing with English composition, (3) course grades in composition classes, (4) grade level in school, and (5) age.

Procedures. To investigate the hypotheses, a Likert-Type Attitudinal Scale was developed to measure attitudes about student-evaluated composition. The Attitudinal Scale consisted of 50 items concerned with the general attitude variables discussed previously, and a final item, (Q51), which was a general statement about how students liked their writing classes.

A sample of 170 students was selected at random from the total population of senior high schools in a Mid-West City. Data collected for the study included scores on the Attitudinal Scale and the *Sequential Test of Educational Progress, Form IA* (STEP), course grades in composition classes, and the demographic variables of age, sex, and grade level in school.

Results. The following general conclusions were drawn from the findings of the study as listed in the sub-variables: (1) The better students, those with A and B grades, welcomed the additional motivation of having one's fellow students evaluate their compositions. (2) The Attitudinal Scale clearly showed that students believed both teachers and students tend to waste time doing the evaluation of each other's composition in writing classes. Further, students expressed a strong interest in having teachers use the extra time created by the student-evaluation process to provide additional help in writing classes. (3) Generally, students believed that working with other students was a beneficial learning experience. (4) The study results on the "Q51" item show that students tend to like their writing classes less than other academic classes. This question, which had some of the lowest overall average scores on the survey, stated that, "In general, students tend to like their writing classes less than other classes."

In addition, there were no statistically significant mean differences among the Total Attitudinal Scores when classified by the demographic variables of sex, age and grade level. However, there were significant differences among the Total Attitudinal Scores when classified by STEP scores and course grades in composition classes.

Summary. The students believed that the process of self-evaluation of compositions could aid in composition learning and could provide some released time for the teacher. The general feeling, however, tended to lessen somewhat as the students moved from the tenth, through the eleventh, to the twelfth grade.

Recommendations. Among the areas suggested by the study as needing further study are the following: (1) Do students improve their writing skills more in self-evaluation classes than they would in regular classes? (2) Why do students dislike their composition classes? (3) Do "C" students only feel negative about their composition classes or about all their studies? (4) How can student-evaluated composition classes help provide additional time for English teachers to assist students with their compositions? (5) And finally, what are the specific reasons for the mean differences being so variable between the course grades (A, B, C)?

THE CONSTRUCTION AND VALIDATION OF THE LISTENING AND READING COMPONENTS OF THE ENGLISH AS A SECOND LANGUAGE ASSESSMENT BATTERY

Order No. 8005254

LOMBARDO, MARIA ELIZABETH, ED.D. *Boston University School of Education*, 1979. 349pp. Major Professor: Maria Estela Brisk

The diagnosis of language proficiency for grouping limited English speaking students in bilingual education programs has been problematic for three reasons: (a) inconsistent identification of language proficiency skills for functioning in a monolingual and/or biligual classroom; (b) non-specification of appropriate norm- and criterion-referenced tests for assessing receptive and expressive English; and (c) limited availability of criteria for grouping bilingual students. To address these problems, the English as a Second Language Assessment Battery (ESLAB), a criterion-referenced measyre was constructed and validated with secondary students.

This study examined the Receptive Area (listening and reading) through the Comprehension Test, the Structural Competency Test, and the Informal Reading Inventory (IRI). The Expressive Area (speaking and writing) was analyzed by Rivera (1979) through the Oral Screening Test, the Oral Competency Test, the Dictation Exercise, and the Writing Sample.

General guidelines for constructing the battery were specified. Then, item analysis, reliability, and validity were established. Teacher training workshops were conducted for the six teachers who examined the students. A sample of 59 mostly Hispanic students from an inner city area was tested. This 12 to 16 year seventh and eighth grade group was composed of 25 boys and 34 girls with varying degrees of English proficiency.

Item analysis based on logical and empirical knowledge included revision and rearrangement of items. The data findings were: (a) percentile (p) values of 27.1% to 89.9% for the Aural Comprehension Test, 1.7% to 64.4% for the Structural Comprehension Test, and 0.0% to 69.5% for the IRI; and (b) the point biserial (RPB) values of .08 to .67 for the Aural Comprehension Test, -.01 to .47 for the Structural Competency Test, and 0.0 to .70 for the IRI.

Reliability involved the computation of internal consistency. The Hoyt estimate of reliability for the Aural Comprehension Test was .81, for the Structural Competency Test .37, and for the IRI .83 and .79 for the total. Cronbach's alpha values were .74 for the IRI and .21 for the total test statistics. The SEm for the three tests was 4.25.

Four types of validity were established: (a) face validity based on examiner's and examinee's judgements, (b) content validity determined by language and reading experts analyzing test items in terms of the specified objectives, (c) predictive validity using Kendall's tau (Nie et. al., 1975) to correlate each test level results (Beginner I, Beginner II, Intermediate I, Intermediate II, and Advanced) with the four Teacher Estimates (T.E.) and ESL report card grades, and (d) concurrent validity used only for the IRI.

A Pearson Coeficient indicated a high correlation between the IRI and Cloze Test .7184 at p .01; IRI and *Stanford Diagnostic Reading Test* (1976 Ed.), .6203 at p .01; Standord and Cloze .6105 at p .05; T.E. and Stanford .4248 at p .05; T.E. and Cloze .3577; and T.E. and IRI - .0456 (Teachers tended to underestimate sutdents' reading levels).

It was concluded that the Receptive Area tests were valid measures of language proficiency. The relationships among the four language areas evidenced from the research were: (a) listening and reading are related; (b) the receptive area is related to the expressive area; and (c) oral skills precede literacy skills.

AN EXPLORATORY STUDY OF DIFFERENCES BETWEEN TWO ENGLISH READING ACHIEVEMENT TESTS WITH HISPANIC BILINGUAL STUDENTS AND MONOLINGUAL ENGLISH STUDENTS

Order No. 8010078

LUBAUS, ZANE, PH.D. *The Pennsylvania State University*, 1979. 218pp. Adviser: Lester S. Golub

The utility of standardized tests normed on and designed mainly for monolingual English speakers has often been questioned for linguistically and culturally different students. This study examined differences in the validity, reliability, and difficulty between comparable levels of two English reading achievement tests with Hispanic bilingual students and with monolingual English students. The two tests examined were the nationally normed Gates-MacGinitie Reading Test and the Inter-American Test of Reading, which is not nationally normed but designed for and used mainly with Hispanic students.

This study also examined (1) the response patterns of Hispanic and non-Hispanic students for evidence of item bias on the Gates-MacGinitie Reading Test and (2) the construct validity of vocabulary and comprehension reading subskills for evidence of convergent and discriminant validity as measured by four different methods.

It was hypothesized that for Hispanic students the Inter-American Test of Reading would have higher validity and reliability, but a lower difficulty level than the Gates-MacGinitie Reading Test. It was further hypothesized that for non-Hispanic monolingual English students the Gates-MacGinitie Reading Test would have higher validity and reliability, but a lower difficulty level than the Inter-American Test of Reading.

The sample population consisted of 252 students in grades four, five, and six from two school districts in southeastern Pennsylvania. Students were divided into the following three language-ethnic groups: Hispanic limited-English-proficiency students from English-Spanish bilingual education classes, Hispanic English-proficient students from monolingual English classes, and non-Hispanic monolingual English students from monolingual English classes.

Validity, reliability, and difficulty indices were established separately for each of the three groups of students for both the Gates-MacGinitie Reading Test, Survey D, Form 1M, and the Inter-American Test of Reading, Level 3, Form CE. Differences between the indices were tested for each group. Validity indices were determined by correlating each of the two tests with each of three criterion-related measures: The Reading Inventory, The Listening Inventory, and The Teacher's Rating Scale of Reading Skills. Reliability indices were obtained by the Kuder-Richardson formula 20 for internal consistency. Difficulty indices were established by computing the mean proportion of items correct on each test, and the response patterns of Hispanic and non-Hispanic students on the Gates-MacGinitie Reading Test were examined. Convergent and discriminant validity patterns of vocabulary and comprehension reading subskills were investigated using multitrait-multimethod matrices.

The findings of the study indicated that for all three groups of students the validity indices slightly favored the Gates-MacGinitie Reading Test over the Inter-American Test of Reading, reliability indices did not differ significantly between the two tests, and the difficulty level of the Gates-MacGinitie Reading Test was lower than of the Inter-American Test of Reading. No evidence of cultural or linguistic bias against Hispanic students was found on 25 items examined on the Gates-MacGinitie Reading Test. Insufficient evidence was found that vocabulary and comprehension subskills of reading are separate and distinct subskills for any of the three groups.

This study provided evidence that the nationally normed Gates-MacGinitie Reading Test appears to be sufficiently valid and reliable for use with Hispanic limited-English-proficiency students and Hispanic English-proficient students. This study also provided evidence that the Gates-MacGinitie Reading Test appears to measure the same traits in similar ways for Hispanic students and for monolingual English students. It was concluded that the Gates-MacGinitie Reading Test may be more appropriate for use with Hispanic students than the Inter-American Test of Reading.

THE RELATIONSHIP BETWEEN ABILITY GROUP PLACEMENT AND TWO STANDARDIZED READING MEASURES, A CRITERION-REFERENCED READING MEASURE, AND A MEASURE OF PARADIGMATIC LANGUAGE AMONG SELECTED FOURTH GRADE SUBJECTS

Order No. 8001027

MILLER, Henrietta Clay, Ph.D. University of Georgia, 1979. 191pp. Major Professor: James H. Dinnan

The purpose of this study was to investigate the relationship between individual fourth grade subjects' actual performance on four selected tests and the placement of these subjects in the school's homogeneous ability levels. The relationship of a subject's sex with his performance on the tests was also considered. More specifically, the purpose of this study was to determine the interrelationships of the following variables: (1) ability level placement; (2) the reading comprehension score on the Iowa Tests of Basic Skills (ITBS); (3) the total reading score on the Metropolitan Achievement Tests (Metropolitan); (4) the total score reading score on the Georgia Criterion-Referenced Tests (Ga. C-RT); (5) the score on the Oral Paradigmatic/Syntagmatic Language Inventory (P/S Inventory); and (6) the sex of the subjects.

The study was designed to investigate the following questions for male subjects, female subjects, and male and female subjects considered together:

1. What is the relationship between the subjects' actual performance on the four measures of reading ability and their ability level placement?
2. What is the relationship of the four measures of reading ability considered together and the students' ability level placement?
3. Can facility in predicting ability level placement be increased significantly by considering a combination of the four variable measures rather than just one?
4. What are the interrelationships between the Ga. C-RT, the ITBS, the Metropolitan, and the P/S Inventory?
5. Does actual performance seem to bear out ability level placement for this grade population? That is, do students in the high ability groups perform better on the achievement measures than students in the lower groups?

The statistical procedures indicated the following:

1. There were significant correlations between performance on the four achievement measures and ability group designation for males, females, and the total sample.
2. For males, females, and the total sample, the combination of all four of the measures considered together was a significantly better predictor of ability group designation than any one of the measures considered alone. The inclusion of each additional measure added significantly to the predictive power of the combination; nevertheless, adding the three other measures to the single measure with the highest correlation accounted for so little additional variance that using all of the combination may not be justified in terms of time, effort, and materials.
3. Scores on the four achievement measures correlated positively with each other for males, females, and the total sample.
4. The rank order of means for scores on the four achievement measures did correspond to the ability group designations. That is, there were significant differences in performance across the ability groups for each of the four achievement measures. The exceptions were groups A and B, from which the Special Education students had been withdrawn. These groups often interchanged in rank. Overall, students who performed best on the achievement measures tended to be in the highest ability groups, and students who performed least well on the achievement measures tended to be in the lowest ability groups.

AN EXAMINATION OF RELATIONAL CONTROL CODING INSTRUMENTATION: CONCEPTUAL, CONSTRUCT, AND PREDICTIVE VALIDITY

Order No. 8007248

MIURA, STEVEN YASUO, PH.D. *Washington State University*, 1979. 117pp. Chairman: H. Joseph Ayres

This study considers the conceptual, construct, and predictive validity of available instruments for coding relational control. Conceptual validity is approached by offering a logical analysis that compares each instrument to the theoretical notions of relational control--multiple communication levels, relationship types, and the interactional perspective--advanced by Gregory Bateson and his associates. Construct validity is approached via an empirical analysis that compares five instruments with each other to determine the extent to which they demonstrate convergent and discriminative strength. Predictive validity is also approached empirically by comparing the coding decisions obtained with these five instruments to the judgments made by native speakers concerning the nature of 29 decision-making interactions.

The results of the logical and empirical analyses indicate that each instrument has different strengths and limitations regarding conceptual, construct, and predictive validity. Explanations for these results are provided and implications for future research in the area of relational control are discussed.

THE CONCURRENT VALIDITY OF CLEP COMPOSITION SCORES IN RELATION TO WRITING PERFORMANCE

Order No. 8007492

OZIER, PAUL WILLIAM, PH.D. *The Florida State University*, 1979. 80pp. Major Professor: Walt Wager

This study investigated the question: What is the relationship between scores achieved on the multiple-choice version of the CLEP Composition General Examination and scores achieved on essays that represent college composition course work? The purpose was to determine the concurrent validity of the CLEP test as an index of writing performance at a single community college through a "case study" approach.

Thirty-three subjects who had taken the CLEP test prepared compositions in a manner similar to common practice in freshman English courses. Each paper was rated by three raters preselected from the English faculty at the participating college. An inter-rater reliability study produced a reliability coefficient of .88. Raw scores for each essay were determined by summing the three ratings for each paper. Pearson r calculations produced a correlation of .58 ($p < .001$) between the test scores and the essay raw scores.

The CLEP Composition General Examination was judged to be slightly successful in indexing writing skill among the sample at the participating community college. It was, however, deemed to be inappropriate as a solitary instrument for granting college credit in composition. A combination of the CLEP test and appropriate writing sample was concluded to be the best means for crediting through examination.

A COMPARATIVE STUDY OF THE PREDICTIVE VALIDITY OF TWO READING READINESS TESTS

Order No. 8001302

PERRY, Leslie Anne Devaney, Ph.D. The University of Mississippi, 1979. 124pp. Director: Associate Professor Charles F. Sherwood

The Problem

The major purpose of this study was to develop and validate the Reading Readiness Inventory. The Inventory was designed by this author to be used by kindergarten and first grade teachers to assess the reading readiness of their students. It consists of 10 subtests: (1) Visual Memory, (2) Recognition of Capital Letters, (3) Recognition of Lower Case Letters, (4) Matching Capital and Lower Case Letters, (5) Matching Word Forms, (6) Rhyming Words, (7) Recognition of Beginning Sounds, (8) Sequencing, (9) Opposites, and (10) Basic Concepts. The individually-administered Inventory takes less than 15 minutes to give and can be administered by an aide or other paraprofessional.

Procedures

During the second week of the 1978-79 school year, the Metropolitan Readiness Test (MRT) was administered to all the first-grade students in the Water Valley, Mississippi Public Schools. The Reading Readiness Inventory was administered to the same students during the third week of school. In April the subjects were given the California Achievement Test (CAT). Complete data from the three tests were available for 117 students. Of these students, 59 were white (34 males, 25 females) and 58 were black (32 males, 26 females). Children who were repeating the first grade were included in the study. Computer programs available through the Statistical Package for the Social Sciences were used to analyze the relationships between the 28 predictor variables and the six criterion variables used in the study. Pearson Coefficients of Correlation were obtained between scores on the two readiness tests and between the readiness test scores and the achievement test scores. These correlations were significant at the .05 level or better. Equations were obtained through multiple regression analysis for predicting reading achievement from the scores on the readiness tests.

Findings

1. In the area of normative information, females outscored males on all the subtests of the Reading Readiness Inventory. On the MRT, females scored higher than males on the visual and language skills subtests but males scored higher on the auditory skills subtests. White students obtained higher scores than black students on all portions of the Reading Readiness Inventory and the MRT.

2. A correlation of .799 was obtained between the total scores on the Reading Readiness Inventory and total pre-reading raw scores on the MRT. A .760 correlation was obtained between the Reading Readiness Inventory total scores and the total reading raw scores on the CAT. This compared with .679 obtained between the MRT and CAT scores.

3. The subtests of the Reading Readiness Inventory which correlated most highly with the total reading raw scores on the CAT were: Matching Capital and Lower Case Letters (.738), Recognition of Lower Case Letters (.717), Recognition of Capital Letters (.689), and Opposites (.634). The subtests of the MRT which correlated most highly with the CAT total reading raw score were Finding Patterns (.633) and Listening (.560).

4. The Pearson Coefficients of Correlation obtained between the demographic variables and the total reading raw scores on the CAT are as follows: Age .005, Teacher .046, Sex .259, and Race .461.

THE USE OF CLOZE PROCEDURE IN THE MEASUREMENT OF SELECTED ENGLISH SYNTACTIC STRUCTURES AND ITS RELATIONSHIP TO COGNITIVE STYLE

Order No. 8010091

REAP, MARY, PH.D. *The Pennsylvania State University*, 1979. 98pp. Adviser: Lester S. Golub

The present study investigated the use of cloze procedure as a means of measuring the difficulty level of selected English syntactic structures found for a group of adult learners of English as a second language. The relationship between scores on a cloze test and cognitive style (field dependence-independence) was also investigated. The purpose of this investigation was: (1) to investigate the degree to which a syntactically derived cloze test can determine the difficulty level of selected English syntactic structures for adult learners of English as a second language, (2) to determine whether the difficulty order found for these syntactic structures varies across language groups, and (3) to investigate the relationship between cognitive style and scores on the syntactically derived cloze test.

One hundred sixty-one subjects from seven language groups were selected from four eastern universities. Subjects completed a cloze test of thirty-seven items. At least three occasions for each of the eight structures were provided on the test. Kendall rank-order correlation coefficients were computed for the difficulty order found among the seven language groups. Significant correlations ($p < .05$) were found for some language groups only.

Inter-sample reliability coefficients were obtained to determine the relationship between the syntactic structures tested on the cloze test and on a multiple-choice test. A statistically significant correlation was obtained between the two measures ($p < .001$).

An independent sample of 124 subjects from the same language groups was given a multiple-choice test of the same English syntactic structures. At least three occasions for each of the eight structures were provided. The correlation between scores on this test and cognitive style was not statistically significant ($p < .05$).

INFORMATION GAIN AS A MEASURE OF READING COMPREHENSION Order No. 8000877

REYNOLDS, Marianne Clifford, Ed.D. Rutgers University The State University of New Jersey (New Brunswick), 1979. 91pp. Chairperson: Edward B. Fry

The purpose of this study was to determine the practicality of using information gain to measure reading comprehension. Two tests were used--a standardized reading comprehension test (California Achievement Test: Reading Comprehension) and a test specifically designed to be passage dependent (Passage Dependent Test). Both tests were administered in an information gain format and a traditional format to one of two groups.

The subjects were 161 eleventh and twelfth graders. Testing took place over a three day period, the students were randomly assigned to one of two testing groups. Raw scores on pre-test, post-test, and traditional test measures were tallied. In addition, an information gain score was computed for each student as the difference between pre- and post-test scores. Student I.Q. scores were made available to the investigator and were correlated with reading scores obtained during the course of the study.

The following questions were examined:

1. Do information gain scores provide us with different information about students' reading comprehension ability from that obtained by traditional reading comprehension measures?
2. Do students gain information from reading passages on comprehension tests?
3. How do the pre-test scores on the California Achievement Test (Reading Comprehension) and the Passage Dependent Test compare with chance scores?
4. What is the relationship between I.Q. scores and traditional reading comprehension scores, and between I.Q. scores and information gain scores?
5. Will a pre-test increase scores on a post-test?
6. What is the relationship between traditional Passage Dependent scores and traditional California Achievement Test (Reading Comprehension) scores?

Analyses of the data obtained in the study were used to answer the questions as follows:

1. Moderate correlations (.45, .46) between gain scores and traditional scores on reading comprehension tests seem to indicate that information gain measures reveal something different or additional about students' reading comprehension.
2. Readers do gain information as a result of reading passages as revealed by the difference between pre- and post-test scores on the CAT and PDT. Readers gained much more information from reading the PDT than they did from the CAT.
3. The comprehension items on the CAT were much less passage dependent than the comprehension items on the PDT as revealed by the differences between a pure chance score and the mean pre-test scores on these two measures.
4. The relationship between I.Q. scores and a standardized traditional reading comprehension test was much higher than the relationship with gain scores. (The PDT used in this study had only a moderate correlation with I.Q. scores even in the traditional form.) The test content and type of question, as well as testing technique, affected the relationship with I.Q.
5. The students who took a pre-test had a slightly, but not significantly, higher mean post-test score than the students who had no pre-test for both measures.
6. A positive relationship between students' PDT and CAT traditional scores was established through a correlation coefficient.

The following conclusions were drawn from the results of this study. Information gain is an alternative technique for measuring reading comprehension with a certain type of testing instrument. It has the advantages of eliminating some of the extraneous factors which affect traditional reading comprehension scores. The pre-test takes into account comprehension items which may be answered correctly on the basis of previous knowledge, test taking skills, etc., rather than on the basis of reading the related passage. Gain scores also have a lower correlation with I.Q. than traditional reading test scores do, an advantage in assessing comprehension as an ability distinct from intelligence.

The Passage Dependent Test is a measurement instrument worthy of consideration for assessing reading comprehension. The advantages are an acceptable reliability coefficient, passage dependent comprehension items, and a limited relationship with I.Q. scores.

THE ORFF TECHNIQUE AS AN INFORMAL DIAGNOSTIC SCREENING FOR READING READINESS Order No. 8012700

RICHARDSON, NELDA SAVACOOL, ED.D. *Northeast Louisiana University*, 1979. 153pp. Major Adviser: Dr. Marjorie S. Snyder

The purpose of this study was to assess the effectiveness of the Orff technique as an Informal Diagnostic Screening of Reading Readiness in kindergarten. The subjects of this longitudinal study were eighty-four students enrolled in Orangewood Elementary School, Fort Myers, Florida. During the investigation activities devised by the researcher applying the Orff technique were used to rate the subjects. Classroom teachers noted selected behaviors during the screening. Following a conference with the teacher and reviewing of a video tape of the screening session, the researcher recorded a score on each student's profile sheet.

Other predictor variables analyzed, along with the Orff technique of Informal Diagnostic Screening of Reading Readiness, were sex, race, socioeconomic status, and scores obtained on the *Look, Listen, and Learn* and the *Stanford Achievement Test* administered in the first and second grades.

Discriminant analysis and canonical correlation were employed in data analysis. Chi-square and Wilks' Lambda were employed and univariate F tests were obtained for each independent variable.

The statistical results seemed to warrant the following conclusions: (1) There appeared to be a relationship among certain ORFF variables (Gross Motor, Auditory, Visual, and Social) and the visual memory subtest from the *Look, Listen, and Learn*. (2) The data reflected a significant correlation between the ORFF Auditory subtest and scores obtained on the *Stanford Achievement Test* in the first grade. (3) There was a positive relationship between the ORFF Visual and Auditory subtests and the *Stanford Achievement Test*, subtest Reading I. (4) The ORFF Visual and Gross Motor subtests also were significant in their discriminating power. (5) There was a 59 percent accuracy on groups formed on the basis of scores on the *Stanford Achievement Test*, the ORFF, sex, race, and socioeconomic status of the subject. (6) The data appeared to indicate a relationship between groups formed on the basis of scores on the *Stanford Achievement Test* and the variables of ORFF, sex, race, and socioeconomic status as discriminant functions in the first grade. The discriminating variables predicted higher percentages of subjects correctly in the highest achieving group and lowest achieving group, which may suggest confounding variables in the middle group. It is suggested that the average group could be studied at a future date and the possibility of I.Q. and absenteeism be investigated as additional variables. (7) ORFF, race, and socioeconomic status attained a significant relationship between groups formed on the basis of *Stanford Achievement Test* scores obtained in the second grade.

A STUDY TO DETERMINE THE OPTIMUM READABILITY LEVEL OF A TEXT IN RELATION TO STUDENTS' READING LEVELS AND ACHIEVEMENT IN A COMMUNITY COLLEGE INTRODUCTORY PSYCHOLOGY COURSE Order No. 8007514

SULLIVAN, VINCENT JOSEPH, ED.D. *The Florida State University*, 1979. 66pp. Major Professor: Maurice L. Litton

The purpose of the study was to determine the optimum readability level of an introductory psychology text in relation to students' reading levels so that community college students would perform satisfactorily on the achievement tests.

The following procedures were completed: (a) administering and scoring achievement tests, and (b) determining the readability level of the text using Flesch's Reading Ease Formula.

To test whether the proportion of achievement scores in two groups was the same, a binomial test of the difference in two proportions was computed.

Results. The hypotheses investigated in the study were: (a) There is no statistically (alpha level = .05) significant increase in the proportion of achievement scores of 75% and above among the students with a reading level two or more years above the readability level of the text compared to students with a reading level at that of the text and up to two years above the text. Hypothesis was rejected. Significantly more students in the group of students reading two or more years above the text made achievement scores of 75% and above than did the group of students reading at the same level of the text and up to two years above the text; (b) There is no statistically significant difference in the proportion of achievement scores of 75% and above among students who are .1 to 1.1 years above and .1 to 1.1 years below the readability level of the text. Hypothesis was not rejected; (c) There is no statistically significant increase in the proportion of achievement scores of 75% and above among the students with a reading level at that of the text and up to two years above the readability level of the text than students with a reading level less than the readability level of the text. Hypothesis was rejected.

Conclusions. Students reading at two or more years above the level of the text had an advantage because they were reading not at the instructional level but at the independent level. The students having reading levels below that of the text had difficulty learning because they were reading at the frustration level. The students having reading levels within and just above that of the text had difficulty learning, even though they were reading at the instructional level, because they experienced frustration in trying to comprehend the text. The college student is not usually able to learn effectively at the instructional level because he has to get most of the material on his own and the textbook is often too difficult to comprehend in this manner.

This study supports Santa and Burstyn's (1977) research which concluded that readability formulas tend to over-estimate the student's ability to comprehend the text. This study goes one step further in suggesting the optimum readability level of an introductory psychology text so that community college students will perform satisfactorily on the achievement tests. This optimum readability level is two or more years below the reading levels of the students because more students perform satisfactorily on achievement tests when they are reading two or more years above the level of the text (independent level).

MOTHERS' ESTIMATIONS OF THEIR CHILDREN'S READING ACHIEVEMENT Order No. 8002145

TRUNNEL, Vesta Mickel, Ed.D. University of Cincinnati, 1979. 62pp.

It is difficult to ascertain which mothers are able to accurately assess their children's achievement and which ones need help. This study extended the study of Vukelich (1978) and investigated the mothers' ability to accurately indicate the reading achievement of their children in grades one through six.

These hypotheses were formulated:

(1) Mothers of children who show high achievement in either Vocabulary or Comprehension sub-tests of the A.S.A.T., will, on average, under-estimate their child's achievement on that sub-test.

(2) Mothers of children who show low achievement in either Vocabulary or Comprehension sub-test of the A.S.A.T., will, on average, over-estimate their child's achievement on that sub-test,

(3) Mothers of children who show average achievement in either Vocabulary or Comprehension sub-test of the A.S.A.T., will, on average, correctly estimate their child's achievement on that sub-test,

(4) At all achievement levels, mothers will more accurately estimate their child's achievement on the Vocabulary section of the A.S.A.T., than on the Comprehension section.

Hypotheses 1, 2 and 4 were supported; hypothesis 3 was not supported.

The subjects of this study were 220 mothers of elementary school students, and the children themselves, from three elementary schools in Northern Kentucky. These mothers were given the American School Achievement Test: (Reading Battery: Vocabulary and Comprehension). They were asked to respond to this test as they thought their children would. The children were given the test the next day at school. The stanine scores of the mothers and students were then compared.

Hypotheses 1, 2 and 3 were tested through a T-test on correlated measures within the child's achievement group for each sub-test. Hypothesis 4 was tested by comparing the mean of mothers' error of prediction at each child's achievement level between the two sub-tests.

The findings of this study showed that mothers were not able to estimate their child's reading achievement at any stanine levels. In examining the accuracy of the mothers' estimations of their child's achievement by grade level of the child, there was a general trend from gross overestimation among mothers of first grade students to marked underestimation among the mothers of sixth grade students, for both sub-tests.

The technique used is applicable at each grade level, but the findings are not consistent at every grade level. The direction of error shifts from grade level to grade level and the pattern of error depends on the grade level where it is used.

THE CLASSIFICATION OF READING COMPREHENSION TEST ITEMS AND ITS RELATION TO THE PERFORMANCE OF SELECTED RACIAL GROUPS Order No. 7928624

WASHINGTON, Elois DeFrantz, Ph.D. The University of Iowa, 1979. 169pp. Supervisor: Associate Professor John McLure

This study reports differences in students' performance on three types of reading comprehension test items -- i.e., textually explicit, textually implicit, and scriptally implicit. A question taxonomy developed by P. David Pearson and Dale B. Johnson (1978) formed the basis for construction of the reading test used in the study. Differences in performance on these three types of reading comprehension questions, and differences between the performances of minority and nonminority students were examined.

The population consisted of 226 high school seniors enrolled in government classes. The students attended three integrated high schools located in two midsize midwestern cities (populations below 350,000).

Comparison of the means on each subscale revealed statistically significant differences in students' performance on each type of reading test item. Means were also scaled according to the relative difficulty of each item type. The mean on the textually explicit items was highest, followed by the textually implicit items and then the scriptally implicit items.

Students were subgrouped according to GPA, class rank, family income, educational expectations, and sex. Performance on each type of test item was again evaluated. The study found that students' average performances on each subscale were hierarchically arranged according to levels of GPA and class rank. Average performances according to educational expectations differed mainly between students who aspired to receive a bachelor's degree or beyond and those who planned on fewer than four years of college education. Students from families with income levels above $20,000 scored significantly better than did students from all income levels below $20,000. Comparison of male and female students' average performances revealed that their mean scores were approximately equal on all subscales except the textually explicit subscale, on which the female students scored significantly better than did the male students.

When comparisons were made between minority and nonminority students' performances on each of the subscales, the result indicated that minority students scored significantly lower than did nonminority students. However, the differences between the scores of the racial groups varied by subscale. The difference between racial groups was smallest on the textually explicit items.

Comparisons were also made between racial groups equated by high and low grade point averages. Differences between the

racial groups' average performances on each of the subscales still existed within the high and low GPA groups. Nonminority students scored significantly higher than did minority students on the total test and the textually implicit and scriptally implicit subscales. However, the difference between the equated GPA groups' performances on the textually explicit subscale was not statistically significant. This was true for both the high and low GPA groups. Evaluation of the average discrimination indexes revealed that the textually explicit items were more effective in discriminating between high- and low-scoring minority students than were either textually implicit or scriptally implicit items. Textually explicit and textually implicit items were equally effective in discriminating between high- and low-scoring nonminority students. Scriptally implicit items discriminated less effectively than either textually explicit or textually implicit items. Nevertheless, scriptally implicit items discriminated significantly better for nonminority students than for minority students.

The study concluded that reliable reading comprehension test items could be constructed to fit the Pearson and Johnson question taxonomy. Furthermore, differences in students' average performance on each of the three types of questions were statistically significant. It was concluded that students' reading performance is probably more accurately tested using only textually explicit and textually implicit items. Inclusion of scriptally implicit items on a reading test would most likely provide less accurate information about minority students' reading abilities than about nonminority students' reading abilities.

A CORRELATIONAL STUDY OF READING PERFORMANCE BETWEEN THE WIDE RANGE ACHIEVEMENT TEST AND TWO FORMS OF THE READING-SPELLING-VOCABULARY PROGRAM PLACEMENT TEST

Order No. 8012729

WHITE, PATRICIA ANNE DAVIS, ED.D. *University of Houston*, 1979. 140pp. Chairman: Dr. Elton G. Stetson

Purpose of the Study. The main purpose of this research study was to investigate the relationship between reading recognition ability on the Stetson READING-SPELLING-VOCABULARY PROGRAM *Placement Test* (RSVP) and the *Wide Range Achievement Test*, reading subtest (WRAT). Specifically, this investigation attempted to: (a) determine the correlational coefficients between performance on a nationally standardized reading test (WRAT) and two forms of the RSVP *Placement Test* (Forms A and B), and (b) complete an item difficulty analysis for each word on the RSVP *Placement Test* (Forms A and B).

Significance of the Study. The review of the literature in this study indicated that there is a need for a current high frequency word list which combines words drawn from children's and adult's writing, spoken language and words in print. To date, there has been no placement test based on such a word list which is small enough to be manageable (1,000 words or less), significant enough to identify the majority of words used in speaking and writing, and broad enough to cover several grade levels. Stetson's READING-SPELLING-VOCABULARY PROGRAM, a synthesis of nine major word lists published from 1913 to 1969, and organized according to grade levels one through nine, is the only graded word list accompanied by a standardized placement test consisting of two equivalent forms to facilitate screening of word recognition ability and accurately place students at the appropriate instructional level in the RSVP.

Procedure and Methodology. The sample population for this research consisted on 269 children selected from grades one through nine in the Cypress-Fairbanks Independent School District and identified to be reading on grade designation as verified by the *Wide Range Achievement Test*. Forms A and B of the RSVP *Placement Test* were administered to the subjects by the investigator and two trained doctoral graduate assistants in May, 1979, within a two week period.

The Pearson Product-Moment Correlation Coefficient was used to determine the relationships between grade scores obtained on the WRAT and the raw scores obtained on both forms of the RSVP *Placement Tests*. Spearman Rank Order Correlation Coefficients were calculated to compare the rank order of words on Form A and Form B of the RSVP *Placement Test* established by Stetson with the rank order determined by the study.

Conclusions. After careful analysis of the data, the following conclusions were warranted: (1) There is a significant positive correlation between scores obtained on the WRAT reading subtest and the RSVP *Placement Test*, Form A. ($r = .861$; $p < .001$); (2) There is significant positive correlation between scores obtained on the WRAT reading subtest and the RSVP *Placement Test*, Form B. ($r = .853$; $p = .001$); (3) There is a significant positive correlation between scores obtained on the RSVP *Placement Test*, Form A and Form B. ($r = .995$; $p = .001$); (4) There is a significant difference in the rank order of the words on each of the subtests of the RSVP *Placement Test* established by Stetson and the rank order determined by the study. (15 out of 16 subtests: rho $< .5139$; df = 13; $p > .05$).

Recommendations. As a result of the findings of the study, several recommendations are suggested: (1) Replication of the study in different school districts in different geographical areas, using a larger sample and allowing for a broader ethnic distribution is suggested. (2) The selection of the sample for replication of the study should not be based solely on reading performance on actual grade designation. (3) Following replications of the study with different populations, data should be analyzed in order to establish grade norms for the RSVP *Placement Test*, Form A and Form B.

Copies of the dissertations may be obtained by addressing your request to:

University Microfilms International
300 North Zeeb Road
Ann Arbor, Michigan 48106

or by telephoning (toll-free) 1-800-521-3042